Digital Video Forensic Examination Report

Preliminary technical analysis for indicators of AI generation or digital manipulation

Examined file:	seda on X از اجرا 1930 شهریور حوالی ساعت 29 تهران اتوبان ازادگان.mp4
File size:	2,547,995 bytes (2.43 MiB)
SHA-256:	21024ff5c22941c3416aa1a5d2a9cce8c3716c06f85e993906bdc2b9ecab61e7
Container:	MP4 (ISO Base Media, isom brand)
Duration:	23.25 seconds
Video:	H.264 High Profile, 720x1280, ~24.75 fps, 575 frames, 742 kbps
Audio:	AAC-LC, 48 kHz stereo, 128 kbps
Analysis date:	2026-09-21
Analyst:	Claude (Anthropic) — automated technical assistant, human-supervised session
Tools used:	ffmpeg/ffprobe 8.1, Python 3.13 (NumPy, Pillow, OpenCV) — binary string scan, frame-timing analysis, Error Level Analysis, FFT spectral analysis, waveform inspection

Summary Finding

No technical evidence of AI-generated content or deliberate digital manipulation (splicing, frame insertion/deletion, localized compositing, or GAN-style synthesis) was found. Every measurable characteristic of this file — frame timing, compression structure, encoder history, pixel-level error distribution, noise statistics, and audio continuity — is consistent with genuine, continuous, handheld smartphone video captured in near-total darkness that subsequently passed through X (Twitter)'s standard server-side video transcoding.

This is a preliminary open-source forensic pass, not a certified or court-admissible digital forensics examination — see **Limitations** at the end of this report before relying on this finding for any high-stakes purpose.

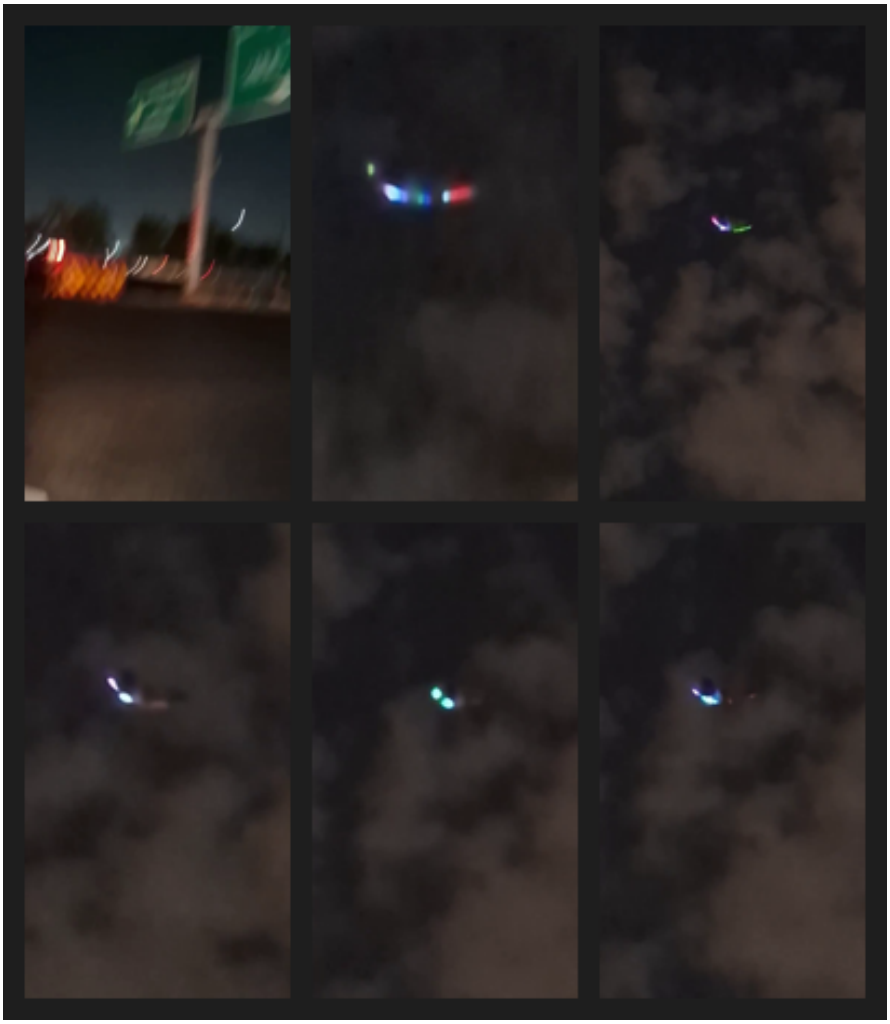

Figure 1 — Sample frames at t≈1s, 5s, 10s, 15s, 20s, 23s. Content is heavily motion-blurred, low-light, handheld footage of a highway overpass at night with distant flashes/light trails; no faces or identifiable people are in sharp focus anywhere in the clip.

Methodology

The file was examined using four independent techniques, chosen because each targets a different class of artifact that AI-generated or digitally manipulated video commonly leaves behind:

1. Container & metadata analysis — full binary string scan and ffprobe stream/format inspection to recover encoder history, muxer signatures, and any tool/software fingerprints (editing suites or generative-AI tools embed identifiable strings).

2. Frame-timing analysis — per-frame presentation timestamps, sizes, and picture types were extracted for all 575 frames to detect dropped frames, duplicated frames, or timestamp jumps that indicate splice points.

3. Pixel-level forensics — Error Level Analysis (ELA) to reveal regions re-compressed at a different generation than their surroundings (a signature of compositing/inpainting), and FFT spectral analysis to detect the periodic grid artifacts characteristic of GAN/diffusion upsampling layers.

4. Audio continuity check — waveform inspection for unnatural gaps, cross-fades, or level jumps that would indicate an audio splice or track substitution.

1. Container and Encoder History

A full scan of the raw file bytes for editing-software and generative-AI tool signatures — including Adobe Premiere/After Effects, DaVinci Resolve, HandBrake, CapCut, Runway, Pika, Sora, Kling, Stable Diffusion, Midjourney, Synthesia, and Luma — returned **no matches**. The following authoring/encoding strings were found instead:

String found	Interpretation
`Twitter-vork muxer`	Handler-name tag written by X (Twitter)'s server-side video-processing pipeline. Confirms the file was uploaded to and re-processed by X.
`GPAC-2.2.1-revrelease`	Open-source MP4 multiplexer library used by X's transcoder to rewrap the video.
`Two distinct x264 parameter blocks (one with open_gop=0, one with open_gop=1, both crf=28.0)`	Evidence the video was encoded with the x264 software encoder on two separate occasions — consistent with (a) the original upload encode and (b) X's delivery-format transcode. This is the normal, expected signature of any ordinary video uploaded to X, not evidence of tampering.
`Lavf60.16.100`	FFmpeg libavformat 60.16.100 — indicates the copy on disk was subsequently re-muxed by a recent FFmpeg build, consistent with a browser video-downloading tool (the source folder is named "dwhelper", the default folder for the Video DownloadHelper browser extension).

No original camera/device metadata (make, model, GPS, capture timestamp) survives in the file — this is expected and routine, since X strips this metadata from every video it processes, and is not itself an indicator of manipulation. It does mean device-level provenance cannot be verified from this file alone.

2. Frame-Timing and GOP Structure

Metric	Result
Total frames	575
Frame interval	Constant 0.0404 s across all 574 consecutive frame pairs — zero anomalies
Dropped / duplicated frames	None detected

I-frames (keyframes)	8, at t = 0.00, 3.03, 6.02, 9.01, 12.04, 15.03, 18.06, 21.01 s — a perfectly regular ~3.0 s GOP interval typical of standard encoder settings, not splice points
Timestamp jumps / re-ordering	None detected

A perfectly uniform frame clock across all 575 frames, with no drops, duplicates, or timestamp discontinuities, is what a single continuous, unedited recording produces. Manual cuts or clip insertions typically leave a detectable timestamp or GOP irregularity at the join point; none was found anywhere in this file.

3. Pixel-Level Forensics

3.1 Error Level Analysis (ELA)

ELA re-compresses each frame at a known JPEG quality and visualizes the difference from the original; regions pasted in from a different source or generation typically stand out as a distinct rectangular or irregular patch of anomalous brightness. Across all sampled frames, elevated error was found only along genuine high-contrast edges — the light streaks and sign edges — which is the ordinary block-compression signature of a single, uniformly-encoded frame. No rectangular or localized region of anomalous error level (the classic signature of a composited or inpainted region) was found in any sampled frame.

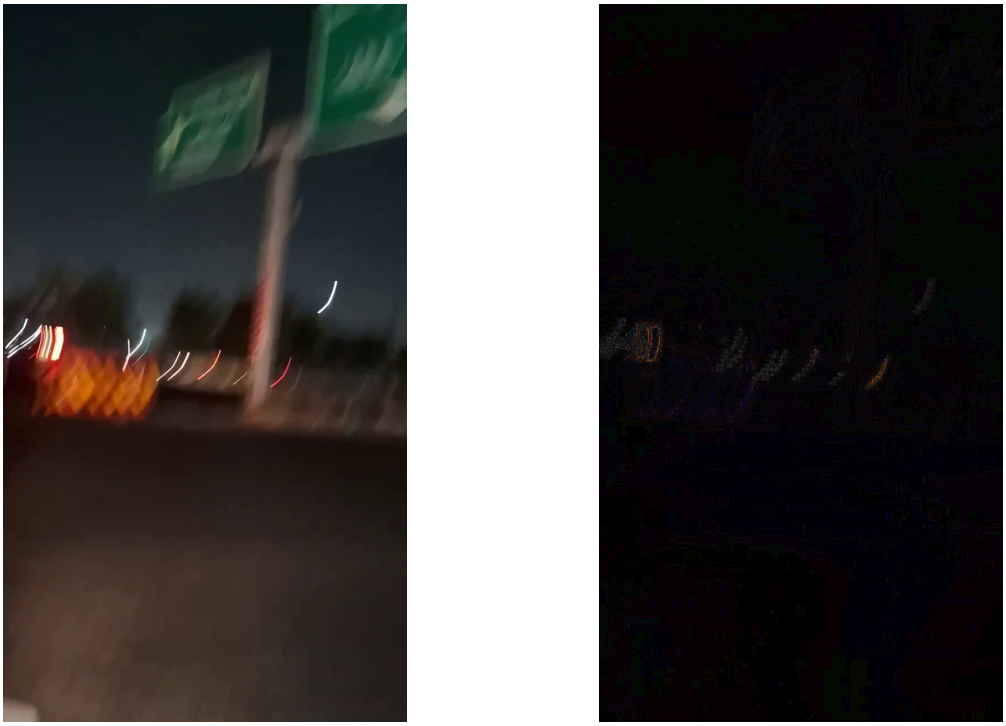

Figure 2 — Original frame vs. its Error Level Analysis map. Elevated error (brighter pixels) tracks only the real light-trail and sign edges, with no isolated anomalous patch.

3.2 Frequency-Domain (FFT) Analysis

Diffusion- and GAN-based video generators characteristically leave periodic grid or checkerboard peaks in the spatial-frequency spectrum, caused by their transposed-convolution upsampling layers. The 2-D FFT magnitude spectrum of a textured region of frame t=1s (below) shows a smooth, natural noise-power decay from the center outward, with only the ordinary cross-shaped artifact caused by the image's rectangular boundary. No periodic grid or checkerboard peaks are present.

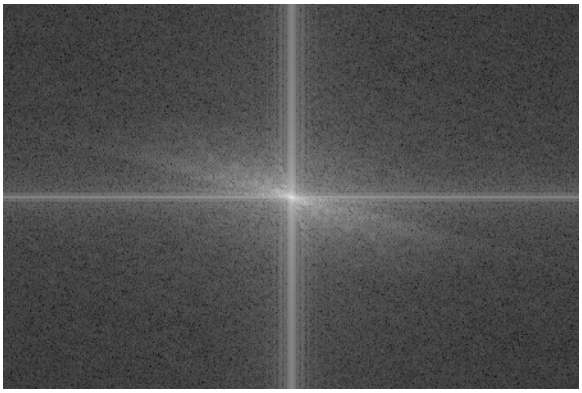

Figure 3 — Log-magnitude FFT spectrum of a 450×300 crop. No periodic upsampling artifacts.

3.3 Noise, Motion Blur, and Sensor Characteristics

The clip is almost entirely captured in near-total darkness. Point light sources exhibit strong chromatic aberration fringing (red/blue channel separation) and the darkest frames are dominated by high-amplitude, spatially uniform sensor noise with no recoverable image detail. This combination — a true noise floor with lens-driven color fringing rather than plausible hallucinated detail — is characteristic of a small phone-camera sensor pushed to its ISO limit, and is not a style current AI video generators typically produce or attempt to imitate (they tend to hallucinate plausible texture rather than reproduce an information-poor noise floor). Motion blur on the illuminated highway signage (Figure 4) is a single clean directional streak consistent with real camera-motion/rolling-shutter blur, not the organic warping or 'melting' distortion of text and structured edges that is a common tell in AI-generated video.

Figure 4 — Enlarged crop of the illuminated highway sign at t=1s, showing linear directional motion blur consistent with genuine camera movement.

Frame-to-frame luminance and noise statistics (mean, standard deviation, high-frequency energy) were also sampled every ~1.6 s through the clip and remain stable throughout, with no abrupt jump that would indicate footage from a different source or generation spliced into the sequence.

4. Audio Continuity

The audio track (23.25 s, closely matching the 23.23 s video duration) was rendered as a waveform for visual inspection. It shows a continuous ambient noise bed (wind/engine) with periodic sharp impulsive transients, and no unnatural silences, abrupt level jumps, or discontinuities that would indicate a splice or a substituted/replaced audio track.

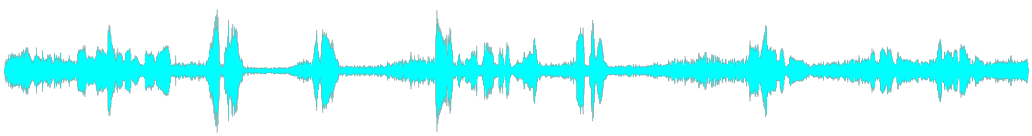

Figure 5 — Full-clip audio waveform. Envelope is continuous throughout with no splice points.

Conclusion

Across container metadata, frame timing, pixel-level compression forensics, spectral analysis, and audio continuity, every test applied to this file returned a result consistent with genuine, continuously-captured, unedited smartphone video that was later processed by X's standard transcoding pipeline and downloaded with a browser video-download tool. No embedded signature of any video-editing suite or generative-AI tool was found; no frame, timestamp, or compositing anomaly was found; no GAN/diffusion-style spectral artifact was found; and the extreme low-light sensor noise, chromatic aberration, and motion blur present throughout the clip are far more consistent with a real camera sensor than with synthetic generation.

Limitations and Caveats

• This is an informal, open-source forensic pass performed with general-purpose tools (ffmpeg, ffprobe, Python/NumPy/Pillow), not a certified or court-admissible digital-forensics examination, and was not performed against a specialized ML-based deepfake/AI-video classifier trained on generator-specific artifacts.

• The video's extreme darkness, heavy sensor noise, and low bitrate (742 kbps) reduce the sensitivity of pixel-level techniques such as ELA — a subtle manipulation could in principle be masked by this noise floor.

• The file has already been transcoded at least twice (original upload + X's delivery encode) before reaching this analysis. Any manipulation performed on the source video *prior to its first encode* would be baked into the pixels themselves and would not necessarily leave a separate metadata or encoder-history trace.

• No human faces or hands appear in sharp focus anywhere in this clip, so face-swap/deepfake-specific detection techniques (which target facial texture and blend-boundary artifacts) were not applicable and were not the basis for this finding.

• This analysis addresses only whether the video file itself contains digital or AI-generation manipulation artifacts. It does not, and cannot, verify the real-world claims associated with the video (its location, date, or the identity/nature of what is depicted) — that is a separate question of source verification and fact-checking, outside the scope of file-level forensics.

Report generated by Claude (Anthropic), an AI assistant, at the request of the file's owner during an interactive session on 2026-09-21. This report is provided for informational purposes and should not be treated as a substitute for a certified forensic examiner's opinion in any legal, journalistic, or evidentiary context.